

WO 00/56897

1

PCI/FR00/00714

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: RHOBIO
- (B) STREET: 14-20 Rue Pierre BAIZET
- (C) CITY: LYONS
- (E) COUNTRY: France
- (F) POSTAL CODE: 69009

(ii) TITLE OF INVENTION: Inducible COMT II promoter, chimeric gene comprising it and transformed plants

(iii) NUMBER OF SEQUENCES: 26

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1863 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 667..672
- (D) OTHER INFORMATION:/function = "inverted W box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 820..830
- (D) OTHER INFORMATION:/function = "inverted L box"

(ix) FEATURE:

- (A) NAME/KEY: enhancer
- (B) LOCATION: 845..852

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1034..1047
- (D) OTHER INFORMATION:/function = "P box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1221..1226
- (D) OTHER INFORMATION:/function = "G box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1343..1356
- (D) OTHER INFORMATION:/function = "inverted
L box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1369..1374
- (D) OTHER INFORMATION:/function = "A box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1377..1382
- (D) OTHER INFORMATION:/function = "GT box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1483..1488
- (D) OTHER INFORMATION:/function = "GT box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1562..1567
- (D) OTHER INFORMATION:/function = "inverted
W box"

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1600..1614
- (D) OTHER INFORMATION:/function = "L box"

(ix) FEATURE:

- (A) NAME/KEY: CAAT_signal
- (B) LOCATION: 1675..1679

(ix) FEATURE:

- (A) NAME/KEY: misc_signal
- (B) LOCATION: 1681..1690
- (D) OTHER INFORMATION:/function = "E box"

(ix) FEATURE:

(A) NAME/KEY: CAAT_signal
(B) LOCATION: 1695..1699

(ix) FEATURE:

(A) NAME/KEY: TATA_signal
(B) LOCATION: 1735..1739

(ix) FEATURE:

(A) NAME/KEY: transcription origin
(B) LOCATION: 1772

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AAAGTTAGGG ACAATCTATA GTGTCACAAA GTTGCTTATG GCTTTTGGTT CAGATAAAGA	60
AAAAGAACAG CATTTTAATT TGTGAAGATT AGTCTGAGCA GAATTTTCATT GTATCTAGAA	120
AGAAATTGAA AAAAGAAATA TTCTATTTCA CTATTATGTT AGGTGCAACT ATATCATCAC	180
CATGGAAAAG CCGGAGTAAA AAGAGAACGT AGAGGAGATT TCATGATTTG ATTGAGAATA	240
TAATATATTA TTTTTTTGTA ATTCCACACA AAGATTAAGA AAATGATCTG ATCAATGATG	300
GCTCCGAGGA TTTGGCTGTC GCGGGAAC TAACATTAAT ATAAATTTGT CGCTGCCTAT	360
AAAGACCC TA TCTATCTATC TATCTATCTA TATATATATA TATATATATA TATATATATA	420
TATATATATA TATATATATA TATATATATA TATATATAAG CGCTAATATT TGATTATTTT	480
TTAAAAATAT TTATAAGTAT ATATGAAATT TTTGACGAAA TTTTGTGTG ACCGTGACCC	540
CTCAACCTAT AGTGTGCGTC CACCTGTGCC AACAAATATAG AGACAATTTG CTCGTATAGT	600

CAGAAAGAGT GTTTACTTT TTAGTTGCTT TTTAGTGAAT CTA CTCTCGGTA TAAAGTAAAA 660
TTAGTGGGTC AATAAGTCGG GTGAATAGTT AAAGAAAACA GTGGTGAGTT TAGCTGTCAA 720
ATAATTTCTT CTTTTCTTG TTTTCACATT AGAAATCAAA ATAAACACA AGCTTTTTGT 780
ATTTATTTTA ACACAAGCTA ATTATATGTT TATATGCTGG TTAGGTGAAG TAAAGCATGT 840
TATATGAGGA AAGTACGAAG AAAATGTGCC AATTGTCGTG TACAGCAAAG CAGCCAGCAC 900
AAGCAAATTC GCACTTGATA AGTGGCTAAG TCCACTTTCT AGTGGACCTA GTGGTTCACT 960
AACTTTTACC AAAAAGGCAA TAATTTGCAA TTCAAAAAGA AAAAAGGAAA AAAGAAAAC 1020
AGACAGACTT TAACACACCA ACTCCACAG GAAGCAACAA TGCAACTCAC AAAAGGAAAC 1080
CGAGTTTTTC CGCGACGGAT CTAGAATTTG GGTTCAATTCT TTACGCTTTT TCGTATTAAA 1140
CTCATTATAT TTGTATAATT ATGGGTTTAT ATTTTTTATT TATTGTAATT TTTGTAAAAT 1200
TTTATATATA AGTGTATACT CCACGTCTCC GGATACTACA TTAGCCTCTA GGGTTCTTAA 1260
TACTCTTGTT AAATTGTCCA GGCTCCAAAC GCATGTTCTG TTCAATTTTA ACGGATGTTT 1320
CCGAACAAC CCAATGTTC AATGTTAGGT GTGTTTGGTG TTAAGCTTCC GTCCTAGGTT 1380
AATAGAATAG ATAATTGTTG TTTCTTATAT AGTTTGAAC AATCGTCGCC ATAACTAAT 1440
TTTTAGGATG GAAGCTAATT TTTAGGATGG AGTACAGCCT AAGGTTAAAA TATAACTATA 1500
AAAAATATCC ATAAAAGGTG AAATTTAATT AGTAACATGA AAAGATAAAA CTAGTGTTAT 1560
CGGTCAAAC TTCAAAAGAG AAAGAAATAA CTAGACAAAC TTCAACAACC AACCTGCCCCA 1620
ACATGCTACT GTGCAATTGA AAAATAAACA AAAGAGAACC AGACAATATT TCAACCAATA 1680
TTCCATCAAG AAAACCAATT ATGACAATTC TTAACCAAAG TCACAACATA CACTTATAAA 1740
AAGCACTAAC TCAACTGTAC ATGATTGTGA AGCCTAACAA AAACACTCTA AAAGGAAAAG 1800
ACTACGAGAA TAATTACACT ACAACTCTTA TAGCTAATTC TTGTCTCAAG ATTTTCAGCT 1860
ATG 1863

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5371 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 1..1860

(ix) FEATURE:

- (A) NAME/KEY: transcription origin
- (B) LOCATION: 1772

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 1861..2281

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 2282..3633

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 3634..3944

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 3945..4726

(ix) FEATURE:

(A) NAME/KEY: exon

(B) LOCATION: 4727..5089

(ix) FEATURE:

(A) NAME/KEY: terminator

(B) LOCATION: 5090..5371

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAAGTTAGGG ACAATCTATA GTGTCACAAA GTTGCTTATG GCTTTTGGTT CAGATAAAGA	60
AAAAGAACAG CATTTTAATT TGTGAAGATT AGTCTGAGCA GAATTCATT GTATCTAGAA	120
AGAAATTGAA AAAAGAAATA TTCTATTTC AATTATGTT AGGTGCAACT ATATCATCAC	180
CATGGAAAAG CCGGAGTAAA AAGAGAACGT AGAGGAGATT TCATGATTG ATTGAGAATA	240
TAATATATTA TTTTTTTGTA ATTCCACACA AAGATTAAGA AAATGATCTG ATCAATGATG	300
GCTCCGAGGA TTTGGCTGTC GCGGGAAC TAACATTAAT ATAAATTTGT CGCTGCCTAT	360
AAAGACCCTA TCTATCTATC TATCTATCTA TATATATATA TATATATATA TATATATATA	420
TATATATATA TATATATATA TATATATATA TATATATAAG CGCTAATATT TGATTATTTT	480
TTAAAAATAT TTATAAGTAT ATATGAAATT TTTGACGAAA TTTTGTGTG ACCGTGACCC	540
CTCAACCTAT AGTGTGCGTC CACCTGTGCC AACAATATAG AGACAATTTG CTCGTATAGT	600
CAGAAAGAGT GTTTTACTTT TTAGTTGCTT TTTAGTGAAT CTAATCGGTA TAAAGTTAAA	660
TTAGTGGGTC AATAAGTCGG GTGAATAGTT AAAGAAAACA GTGGTGAGTT TAGCTGTCAA	720
ATAATTTCTT CTTTTCTTG TTTTCACATT AGAAATCAAA ATAAAACACA AGCTTTTGT	780
ATTTATTTTA ACACAAGCTA ATTATATGTT TATATGCTGG TTAGGTGAAG TAAAGCATGT	840
TATATGAGGA AAGTACGAAG AAAATGTGCC AATTGTCTG TACAGCAAAG CAGCCAGCAC	900
AAGCAAATTC GCACTTGATA AGTGGCTAAG TCCACTTTCT AGTGGACCTA GTGGTTCACT	960
AACTTTTACC AAAAAGGCAA TAATTTGCAA TTCAAAAAGA AAAAAGGAAA AAAGAAAAC	1020
AGACAGACTT TAACACACCA ACTCCCACAG GAAGCAACAA TGCAACTCAC AAAAGGAAAC	1080
CGAGTTTTTC CGCGACGGAT CTAGAATTTG GGTTCATTCT TTACGCTTTT TCGTATTAAA	1140
CTCATTATAT TTGTATAATT ATGGGTTTAT ATTTTTTATT TATTGTAATT TTTGTAAAT	1200
TTTATATATA AGTGTATACT CCACGTCTCC GGATACTACA TTAGCCTCTA GGGTTCTTAA	1260

TACTCTTGTT	AAATTGTCCA	GGCTCCAAAC	GCATGTTTCGT	TTCAATTTTA	ACGGATGTTT	1320
CCGAACAAC	CCAAATGTTT	AATGTTAGGT	GTGTTTGGTG	TTAAGCTTCC	GTCCTAGGTT	1380
AATAGAATAG	ATAATTGTTG	TTTCTTATAT	AGTTTTGAAC	AATCGTCGCC	ATAAACTAAT	1440
TTTTAGGATG	GAAGCTAATT	TTTAGGATGG	AGTACAGCCT	AAGGTAAAA	TATAACTATA	1500
AAAAATATCC	ATAAAAGGTG	AAATTTAATT	AGTAACATGA	AAAGATAAAA	CTAGTGTTAT	1560
CGGTCAAAC	TTCAAAAGAG	AAAGAAATAA	CTAGACAAAC	TTCAACAACC	AACCTGCCCA	1620
ACATGCTACT	GTGCAATTGA	AAAATAAACA	AAAGAGAACC	AGACAATATT	TCAACCAATA	1680
TTCCATCAAG	AAAACCAATT	ATGACAATTC	TTAACCAAG	TCACAACATA	CACCTATAAA	1740
AAGCACTAAC	TCAACTGTAC	ATGATTGTGA	AGCCTAACAA	AAACACTCTA	AAAGGAAAAG	1800
ACTACGAGAA	TAATTACACT	ACAACTCTTA	TAGCTAATTC	TTGTCTCAAG	ATTTTCAGCT	1860
ATGGAATCCT	CAACCAAAAG	CCAAATACCA	ACACAATCAG	AAGAAGAGCG	TAAGTGCACA	1920
TATGCCATGC	AACCTATTGC	ATCTTCAGTC	CTCCCTTTG	TGTTGCATTC	AACAATTCAA	1980
TTGGAAGTTT	TTGAGATATT	AGCCAAATCT	AATGACACTA	AACCTTCTGC	TTCTCAAATT	2040
GTTTCTCAAA	TTCCTAACTG	CACAAAACCT	GAAGCACCTA	CTATGTTAAA	TAGGATGCTT	2100
TATGTCTTGG	CTAGTTACTC	CTTGTTTACT	TGTTCATTG	TTGAAGATGA	AAAAAATAAT	2160
GGGGGCCAAA	AAAGAGTGTA	TGGTTTGTC	CAAGTGGGAA	AATTCTTTGT	TAAAAATGAA	2220
AATGGTGCAT	CAATGGGGCC	ACTTTTGGCT	TTGCTTCAA	ATAAAGTATT	CATAAACAGC	2280
TGGTAAGTTT	TGTCCTACTG	TGTATTCTTT	TTGCACTGGC	TGTATTGATT	GGTTGCCTTT	2340
TTCAACAAGAC	AAGATTCTTA	AGTTTTATTA	CTTGTCGATT	TATGTTAGTC	GTATGTGCTA	2400
GTGTTATTAT	TCTCCATCTG	ATCCTTTTAT	TGGTCACTTT	ACCTAAAAAT	ATTGTTACAA	2460
AACATTTGTC	CTTCTAGAAA	ATCAGGTATT	ATTAATTTTT	CAATTCCATC	TTTATTACTC	2520
CAATAGTGAA	TATGGTTATT	AATTAGTGTT	TTAAGGAAGA	TGTAAGGATA	ATTTAATCAA	2580
ATAGGATTTA	TTATTAATGT	TGTCAAAGAT	TCTGGTGGAT	GGATCGGAGA	AAATTTCTTC	2640
ATCTTAATCA	GAGTTTGATG	TTGAGCCAC	AGGAATGAAT	TTGTTTTTAA	TAGGGAGTAT	2700
TTTCTCTTTG	AATAGACCTT	ACACAATAAA	AGGACAACCC	GGTACACTAA	GCTTCCGTTA	2760
TGCGCGGGGT	TCGGGAAAG	GACCGCATCA	CCAGGTCTAT	TGTACGCAGC	GTTACCCAAC	2820
GTGAATCTAA	ATTAATGAGA	CTAAAAAATG	GAACCAACA	CCAGTGAAAA	CCAAAAAAG	2880
AAGCAAACCT	TAGTGGATGG	CTTGGAAGA	TCTTTCTTCT	TGAATAACTT	GGAGCGCTAT	2940
ATATTAAGGC	GTCGCAGCCG	TTAGATACTT	TCAAGAAGAA	AGCTAAAAAA	TGTTTTAAAG	3000
TTACGGCGCT	AGAATAATGA	AATTTCTCTA	TATATATAAT	TCAAAAGTTA	ATAATTTATT	3060
CTCTTAACCT	AAATCTATAT	TATAAACTA	TATTAAGTAA	CTTCTGCCTA	ATTTATAATA	3120
TACAACTAAT	GTTTTGAGAA	AACAAAATAA	CAACAACATC	AAACCAATG	AAATCCCACA	3180
AGTAGAGTTT	GGGAGGATA	GTGTGTACGG	AGACCTTACC	CCTACCTTAT	AAAGTTAAAG	3240

AGGCTGTTTT CGAAAGACTC TCGGCTCAAG AACATTAAAA ATTTGAGAAA ACAAATATA 3300
AATTCAAAAC CTATATTAAG TTTATAATCC ATGGTATATT ATATTGGCTT AGTAATCTGA 3360
AATGAAAGAT TTATGTTTGA CTCCTCTAAA CTTGTTTTTA ATGCAAAAGA GGCACAACAT 3420
ATATATTATA AGTATCTTTT TTTGGTTTCC CACTGTGGCC GCTAAATTCG GATTTCGCTGG 3480
AAGTGTCCACA TTGTTGGAGA TGGGGGCAAC GCTCACAACA AAGACGATTC TATAATTAGT 3540
GTTCCAACCT GAAATTTTAG TTAAAGATAA AGAAGTACTT ACCATAATGG TAGATATGAT 3600
CA7ATCTGAC TCTCTTTCTA ATTTCAAATT ACAGGTTTGA ACTAAAAGAT GCAGTTCTTG 3660
AAGGAGGAGT TCCATTTGAC AGGGTACACG GTGTGCATGC ATTTGAATAT CCAAAATCGG 3720
ACCCAAAATT CAATGATGTT TTCAACAAGG CAATGATCAA TCACACAACT GTAGTCATGA 3780
AAAAAATACT TGAATAATTAC AAAGGTTTTG AGAACCTTAA AACTTTGGTT GATGTTGGAG 3840
GTGGTCTTGG AGTTAACCTC AAGATGATTA CATCTAAATA CCCACAATT AAGGGCACTA 3900
ATTTTGATTT GCCACATGTT GTTCAACATG CCCCTTCCTA TCCTGGTACC TTCTCTCGTT 3960
CTTATTTTGT TGTATTATTAT ATTTACTTCG ATCATCAGGT CTAGGTCTGT CAAGTTAAAT 4020
TCGTTCTCAA AAAAGTTTAT AAAGGTTTTG AACTCCATCA CCTATTGCTT TAGGATTTTG 4080
AGTTGTATGC TCTGAGTCTT GCGCATGGTA TCATAGTCAA TTTATTTAAG CTCGTTATTG 4140
CACTTGTGAA TTCTATTATA TAAGGAGTAA GCCTACCAA AAGGAGCGAA AATATTTTCC 4200
AAAACCTCTT TTAAACCTTC CTCACCCCAT TCCCCTCTCC CCTCTCCCC AACACCACCC 4260
ACCACCCCAA CTCCTCCGTC TTAGTTTTTT TATTTATCCT GGACTTTCTT ATATTTTATG 4320
CTTTCCTTTA ATTGAACTCT TGTAATAAA CCATTTGCC CCCACCCTAT AGTGTTTGCC 4380
TAAATTTTAT ATTTTTCAAA ATAATATTTT CTATTTACTA ATTAAACATT AGAAAATATT 4440
TTTCGGATTT TTTTCCACTC ACCAACCAAG CATGSGAAAA TAGTGATAAA ACTACTCATT 4500
TTTCAAAATA ATATTTTCAA GGAAAACATT TTCCTTTATA CCAAATACCC TTACTCTTGT 4560
ATACAAATCT TCATGTCGAT GATCTTGCAA TATATATACA TGTATATGTA TGATTTGATA 4620
AACCACATGA ACAAATGGT TGAGCTCTGC GAATTGTGAT ATATGATTTG CTTATGTGTT 4680
GTGCACTATC AATTACTTAA ATTAACTTC ATCTAATAAT ATTGCAGGGG TGGAACATGT 4740
TGGGGGAGAT ATGTTTGAAA GTGTTCCAGA AGGAGATGCT ATTTTATGA AGTGGATTCT 4800
TCATGACTGG AGTGATAGTC ACAACCTCAA GTTGCTAAAG AACTGCTACA AGGCTCTACC 4860
AGACAATGGA AAGGTGATTG TTGTTGAGGC CATTTTACCA GTGAAACCAG ACATTGACAC 4920
CGCAGTGGTT GCGTTTCGC AATGTGATTT GATCATGATG GCTCAAAATC CTGGAGGCAA 4980
AGAGCGATCG GAAGAGGAGT TTCGAGCCTT GGCTACTGAA GCTGGATTCA AAGGCGTTAA 5040
CTTAATATGT TGTGTCTGTA ATTTTGGGT CATGGAATTC TGCAAGTAGA TTTCTACTGT 5100
ACATTGAGTT TCTACTACTC TTGAGTATCC ATTTATGGCA ATCTGGGACT GGAATTGCAG 5160
CTTAGTCCAG ATTGAACATT GATATTCCTA ATAATATTTT TATTATTTCC CTTGTTTATT 5220

TCTCTGTAT GAAAGGATGT CATTTTGAGT ATTGATAATC ATGTTCTCTA GGACAGAAAT 5280
TGTAACCTTG TCCAACCTTA TTGATATTCC TAGTAAGATT TATATGACAT GTGTCTCTGG 5340
TTTGAGAAGA GTTCAATAT CTACAGACGG G 5371

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1095 base pairs
 (B) TYPE: nucleotide
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..1095

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATG GAA TCC TCA ACC AAA AGC CAA ATA CCA ACA CAA TCA GAA GAA GAG	48
Met Glu Ser Ser Thr Lys Ser Gln Ile Pro Thr Gln Ser Glu Glu Glu	
1 5 10 15	
CGT AAC TGC ACA TAT GCC ATG CAA CTA TTG TCA TCT TCA GTC CTC CCC	96
Arg Asn Cys Thr Tyr Ala Met Gln Leu Leu Ser Ser Ser Val Leu Pro	
20 25 30	
TTT GTG TTG CAT TCA ACA ATT CAA TTG GAA GTT TTT GAG ATA TTA GCC	144
Phe Val Leu His Ser Thr Ile Gln Leu Glu Val Phe Glu Ile Leu Ala	
35 40 45	
AAA TCT AAT GAC ACT AAA CTT TCT GCT TCT CAA ATT GTT TCT CAA ATT	192
Lys Ser Asn Asp Thr Lys Leu Ser Ala Ser Gln Ile Val Ser Gln Ile	
50 55 60	
CCT AAC TGC ACA AAA CCT GAA GCA CCT ACT ATG TTA AAT AGG ATG CTT	240
Pro Asn Cys Thr Lys Pro Glu Ala Pro Thr Met Leu Asn Arg Met Leu	
65 70 75 80	
TAT GTC TTG GCT AGT TAC TCC TTG TTT ACT TGT TCC ATT GTT GAA GAT	288
Tyr Val Leu Ala Ser Tyr Ser Leu Phe Thr Cys Ser Ile Val Glu Asp	
85 90 95	
GAA AAA AAT AAT GGG GGC CAA AAA AGA GTG TAT GGT TTG TCA CAA GTG	336
Glu Lys Asn Asn Gly Gly Gln Lys Arg Val Tyr Gly Leu Ser Gln Val	
100 105 110	
GGA AAA TTC TTT GTT AAA AAT GAA AAT GGT GCA TCA ATG GGG CCA CTT	384
Gly Lys Phe Phe Val Lys Asn Glu Asn Gly Ala Ser Met Gly Pro Leu	
115 120 125	
TTG GCT TTG CTT CAA AAT AAA GTA TTC ATA AAC AGC TGG TTT GAA CTA	432
Leu Ala Leu Leu Gln Asn Lys Val Phe Ile Asn Ser Trp Phe Glu Leu	
130 135 140	
AAA GAT GCA GTT CTT GAA GGA GGA GTT CCA TTT GAC AGG GTA CAC GGT	480
Lys Asp Ala Val Leu Glu Gly Gly Val Pro Phe Asp Arg Val His Gly	
145 150 155 160	
GTG CAT GCA TTT GAA TAT CCA AAA TCG GAC CCA AAA TTC AAT GAT GTT	528
Val His Ala Phe Glu Tyr Pro Lys Ser Asp Pro Lys Phe Asn Asp Val	
165 170 175	

TTC AAC AAG GCA ATG ATC AAT CAC ACA ACT GTA GTC ATG AAA AAA ATA	576
Phe Asn Lys Ala Met Ile Asn His Thr Thr Val Val Met Lys Lys Ile	
180 185 190	
CTT GAA AAT TAC AAA GGT TTT GAG AAC CTT AAA ACT TTG GTT GAT GTT	624
Leu Glu Asn Tyr Lys Gly Phe Glu Asn Leu Lys Thr Leu Val Asp Val	
195 200 205	
GGA GGT GGT CTT GGA GTT AAC CTC AAG ATG ATT ACA TCT AAA TAC CCC	672
Gly Gly Gly Leu Gly Val Asn Leu Lys Met Ile Thr Ser Lys Tyr Pro	
210 215 220	
ACA ATT AAG GGC ACT AAT TTT GAT TTG CCA CAT GTT GTT CAA CAT GCC	720
Thr Ile Lys Gly Thr Asn Phe Asp Leu Pro His Val Val Gln His Ala	
225 230 235 240	
CCT TCC TAT CCT GGG GTG GAA CAT GTT GGG GGA GAT ATG TTT GAA AGT	768
Pro Ser Tyr Pro Gly Val Glu His Val Gly Gly Asp Met Phe Glu Ser	
245 250 255	
GTT CCA GAA GGA GAT GCT ATT TTT ATG AAG TGG ATT CTT CAT GAC TGG	816
Val Pro Glu Gly Asp Ala Ile Phe Met Lys Trp Ile Leu His Asp Trp	
260 265 270	
AGT GAT AGT CAC AAC CTC AAG TTG CTA AAG AAC TGC TAC AAG GCT CTA	864
Ser Asp Ser His Asn Leu Lys Leu Leu Lys Asn Cys Tyr Lys Ala Leu	
275 280 285	
CCA GAC AAT GGA AAG GTG ATT GTT GTT GAG GCC ATT TTA CCA GTG AAA	912
Pro Asp Asn Gly Lys Val Ile Val Val Glu Ala Ile Leu Pro Val Lys	
290 295 300	
CCA GAC ATT GAC ACC GCA GTG GTT GGC GTT TCG CAA TGT GAT TTG ATC	960
Pro Asp Ile Asp Thr Ala Val Val Gly Val Ser Gln Cys Asp Leu Ile	
305 310 315 320	
ATG ATG GCT CAA AAT CCT GGA GGC AAA GAG CGA TCG GAA GAG GAG TTT	1008
Met Met Ala Gln Asn Pro Gly Gly Lys Glu Arg Ser Glu Glu Glu Phe	
325 330 335	
CGA GCC TTG GCT ACT GAA GCT GGA TTC AAA GGC GTT AAC TTA ATA TGT	1056
Arg Ala Leu Ala Thr Glu Ala Gly Phe Lys Gly Val Asn Leu Ile Cys	
340 345 350	
TGT GTC TGT AAT TTT TGG GTC ATG GAA TTC TGC AAG TAG	1095
Cys Val Cys Asn Phe Trp Val Met Glu Phe Cys Lys	
355 360	

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide No.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CGTTTCGCAA TGTGATTGA TC

22

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide No.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTCAAAATGA CATCCTTTCA TAC

23

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide No.3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTGAAGATGT CAATAGTTGC ATGGC

25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PAS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGTCTAGAGG GCCTTTTAGA GTGTTTTTGT TAG

33

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAAGTCGACC GTCCACCTGT GCCAACAAAT

29

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGTTTGGTGT TATGCTTCCG TCCT

24

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 292 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AAAAAGCTTT TTTAGGATGG AGTACAGCC

29

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TTTAAGCTTA AAGAGAACCA GACAATATT

29

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 354 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..60
- (D) OTHER INFORMATION:/function = preprotein

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 61..60
- (D) OTHER INFORMATION:/function = preprotein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```

atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc gga 48
Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val Gly
1 5 10 15

tct gcc aac gcc acc gcg tgc acc gcc acc cag cag acc gct gcg tac 96
Ser Ala Asn Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr
20 25 30

aag aca ctc gtg agc atc ctg tgc gac gcg tgc ttc aac aag tgc tct 144
Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser
35 40 45

acg gat tgc ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg 192
Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr
50 55 60

gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg atc 240
Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile
65 70 75 80

aag aag atc gtg acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc 288
Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro
85 90 95

acg agc ggc ctg gtg ctc aac gtg tac tgc tac gcg aac ggc ttc tgc 336
Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser
100 105 110

gac aag tgc tgc tgc ctg 354
Asp Lys Cys Ser Ser Leu
115

```

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```

acc gcg tgc acc gcc acc cag cag acc gct gcg tac aag aca ctc gtg 48
Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala Tyr Lys Thr Leu Val
  1           5           10           15

agc atc ctg tcg gac gcg tcg ttc aac aag tgc tct acg gat tcg ggc 96
Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys Ser Thr Asp Ser Gly
          20           25           30

tac tcc atg ctg acg gcc aag gcc ctc ccc acc acg gcg cag tac aag 144
Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr Thr Ala Gln Tyr Lys
          35           40           45

ctc atg tgc gcg tcc acg gca tgc aac acc atg atc aag aag atc gtg 192
Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met Ile Lys Lys Ile Val
          50           55           60

acg ctg aac ccg ccc aac tgc gac ctg acg gtg ccc acg agc ggc ctg 240
Thr Leu Asn Pro Pro Asn Cys Asp Leu Thr Val Pro Thr Ser Gly Leu
          65           70           75           80

gtg ctc aac gtg tac tcg tac gcg aac ggc ttc tcg gac aag tgc tcg 288
Val Leu Asn Val Tyr Ser Tyr Ala Asn Gly Phe Ser Asp Lys Cys Ser
          85           90           95

tcg ctg 294
Ser Leu

```

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1620 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: COMT II promoter
- (B) LOCATION: 1..1263

(ix) FEATURE:

- (A) NAME/KEY: CDS megaspermine
- (B) LOCATION: 1264..1630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```

cgccacctg tgccaacaat atagagacaa ttgctcgtg tagtcagaaa gagtgtttta 60
cttttttagtt gctttttagt gaatctactc ggtataaagt taaatttagtg ggtcaataag 120
tcgggtgaat agttaagaa aacagtgggtg agtttagctg tcaaataatt tcttcttttt 180
cttgttttica cattagaaat caaaataaaa cacaagcttt ttgtatttat tttacacaa 240
gctaattata tgtttatatg ctggtttaggt gaagtaaagc atgttatatg aggaaagtac 300
gaagaaaatg tgccaattgt cgtgtacagc aaagcagcca gcacaagcaa attcgcactt 360
gataagtggc taagtccact ttctagtggg cctagtgggt cactaacttt taccaaaaag 420
gcaataattt gcaattcaaa aagaaaaaag gaaaaaagaa aactagacag actttaacac 480
accaactccc acaggaagca acaatgcaac tcacaaaagg aaaccgagtt tttccgcgac 540
ggatctagaa ttggggttca ttctttacgc tttttcgtat taaactcatt atatttgtat 600
aattatgggt ttatattttt tattttattgt aatttttcta aaattttata tataagtgt 660
tactccacgt ctccggatac tacattagcc tctaggggtc ttaatactct tgtaaattg 720
tccaggctcc aaacgcatgt tcgtttcaat tttaacggat gtttccgaac aactccaaat 780
gttcaatgtt aggtgtgttl ggtgttaagc ttccgtccta ggtaaataga atagataatt 840
gttgtttctt atatagtttt gaacaatcgt cgccataaac taatttttag gatggaagct 900
aatttttagg atggagtaca gcctaagggt aaaatataac tataaaaaat atccataaaa 960
ggtgaaattt aattagtaac atgaaaagat aaaactagtg ttatcgggtc aactttcaaa 1020
agagaaaagaa ataactagac aaacttcaac aaccaacctg cccaacatgc tactgtgcaa 1080
ttgaaaaata aacaaaagag aaccagacaa tatttcaacc aatattccat caagaaaacc 1140
aattatgaca attcttaacc aaagtcacaa ctaacactta taaaaagcac taactcaact 1200
gtacatgatt gtgaagccta acaaaaacac tctaaaaggc ctctagagga tccccggggt 1260

acc atg aac ttc acc gct ctg ctc gct gcc gtc gcc gcc gcc ttg gtc 1308
Met Asn Phe Thr Ala Leu Leu Ala Ala Val Ala Ala Ala Leu Val
1 5 10 15

gga tct gcc aac gcc acc gcg tgc acc gcc acc cag caa acc gct gcg 1356
Gly Ser Ala Asn Ala Thr Ala Cys Thr Ala Thr Gln Gln Thr Ala Ala
20 25 30

tac aaa aca ctc gtg agc atc ctg tcg gac gcg tcg ttc aac aag tgc 1404
Tyr Lys Thr Leu Val Ser Ile Leu Ser Asp Ala Ser Phe Asn Lys Cys
35 40 45

tct acg gat tcg ggc tac tcc atg ctg acg gcc aag gcc ctc ccc acc 1452
Ser Thr Asp Ser Gly Tyr Ser Met Leu Thr Ala Lys Ala Leu Pro Thr
50 55 60

acg gcg cag tac aag ctc atg tgc gcg tcc acg gca tgc aac acc atg 1500
Thr Ala Gln Tyr Lys Leu Met Cys Ala Ser Thr Ala Cys Asn Thr Met
65 70 75

atc aaa aaa atc gtg acg ctg aac ccg ccc aac tgc aac ctg acg gtg 1548
Ile Lys Lys Ile Val Thr Leu Asn Pro Pro Asn Cys Asn Leu Thr Val
80 85 90 95

ccc acg agc ggc ctg gtg ctc aac gtg tac tcg tac cca aac ggc ttc 1596
Pro Thr Ser Gly Leu Val Leu Asn Val Tyr Ser Tyr Pro Asn Gly Phe
100 105 110

tcg gac aag tgc tcg tcg ctg taa 1620
Ser Asp Lys Cys Ser Ser Leu
115

```

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PAS2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGCGGATCCC CTTTtagagt GTTTTGTTA GGC

33

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ACGCGTCGAC GTTAGGGACA ATCTATAGTG TCAC

33

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ACGCGTCGAC GCTCCGAGGA TTTGGCTGTC GCGG

34

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

ACGCGTCGAC GCTGGTTAGG TGAAGTAAAG CATG

34

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ACGCGTCGAC GCATGTTATA TGAGGAAAGT ACG

33

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACGCGTCGAC GCAGCCAGCA CAAGCAAATT CGC

33

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ACGCGTCGAC GACTTTAACA CACCAACTCC C

31

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ACGCGTCGAC CGGATCTAGA ATTTGGGTTC ATTC

34

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ACGCGTCGAC GTGTATACTC CACGTCTCCG GATAC

35

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

ACGCGTCGAC GTTCAATGTT AGGTGTGTTT GG

32

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PAS3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CGCGGATCCG CTTAACACCA AACACACCTA ACATTG

36

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Synthetic oligonucleotide PS14

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ACGCGTCGAC CAGTGGTGAG TTTAGCTGTC

30